

AD-A207 575

THE H-P VERSION OF THE FINITE ELEMENT METHOD FOR
PARABOLIC EQUATIONS PART.. (U) MARYLAND UNIV COLLEGE
PARK DEPT OF MATHEMATICS I BABUSKA ET AL. FEB 89

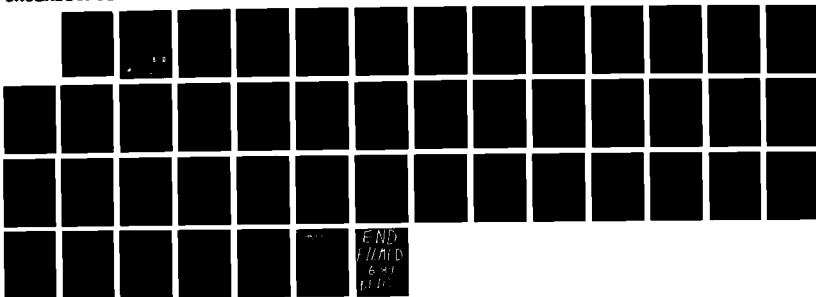
1/1

UNCLASSIFIED

AD89-04-1B-TJ N00014-85-K-0169

F/G 12/1

NL



AD-A207 575



COLLEGE PARK CAMPUS

THE h-p VERSION OF THE FINITE ELEMENT METHOD
FOR PARABOLIC EQUATIONS

PART II: The h-p Version in Time

I. Babuška

Institute for Physical Science and Technology
University of Maryland, College Park, MD 20742

and

Tadeusz Janik

Department of Mathematics
University of Maryland, College Park, MD 20742

MD89-04-IB-TJ

TR89-04

BN-1093

February 1989

DTIC
ELECTE
MAY 10 1989
S E D
Cb



INSTITUTE FOR PHYSICAL SCIENCE
AND TECHNOLOGY

This document has been approved
for public release and sale; its
distribution is unlimited.

89 5 10 031

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER BN-1093	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The h-p Version of the Finite Element Method for Parabolic Equations Part II: The h-p Version in Time		5. TYPE OF REPORT & PERIOD COVERED Final life of the contract
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) I. Babuska ¹ and T. Janik		8. CONTRACT OR GRANT NUMBER(s) ONR N00014-85-K-0169 NSF DMS-8516191
9. PERFORMING ORGANIZATION NAME AND ADDRESS Institute for Physical Science and Technology University of Maryland College Park, MD 20742		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Department of the Navy Office of Naval Research Arlington, VA 22237		12. REPORT DATE February 1989
		13. NUMBER OF PAGES 41
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release: distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The paper is the second in the series addressing the h-p version of the finite element method for parabolic equations. The present paper addresses the case when in both variables, the spatial and time, the h-p version is used. Error estimation is given and numerical computations are presented.		

THE h-p VERSION OF THE FINITE ELEMENT METHOD
FOR PARABOLIC EQUATIONS

PART II - The h-p Version in Time

Ivo Babuška
Institute for Physical Science and Technology
University of Maryland
College Park, MD 20742

and

Tadeusz Janik
Department of Mathematics
University of Maryland
College Park, MD 20742

MD89-04-IB-TJ

TR89-04

BN-1093

February 1989

Partially supported by the ONR Contract N00014-85-K-0169 and NSF Grant
DMS-8516191.

Abstract. The paper is the second in the series addressing the h-p version of the finite element method for parabolic equations. The present paper addresses the case when in both variables, the spatial and time, the h-p version is used. Error estimation is given and numerical computations are presented.

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



1. Introduction.

The paper is the second in the series about the h-p version of the finite element method for solving parabolic partial differential equations.

In the first paper [1] we discussed the case when in the t direction only one element of degree $q \rightarrow \infty$ was used.

In this paper we analyze the case when in both variables, the spatial and time, the h-p version is used. We are showing some essential differences between the p and h-p version. We will keep the notation of [1], but it is not necessary to prerequisite for the present paper.

2.1. The h-p version for the initial value problem for an ordinary differential equation.

Let $I = (0, T)$, $\bar{I} = [0, T]$, $t \in I$, $X = L_2(I)$ be the usual space with the norm

$$(2.1.1) \quad \|u\|_X = \left(\int_0^T u^2 dt \right)^{1/2}.$$

Let

$$\dot{C}^1 = \{v \in C^\infty(\bar{I}) \mid v(T) = 0\},$$

where $C^\infty(\bar{I})$ is the space of functions with all continuous derivatives on \bar{I} .

For any $\lambda > 0$ and $v \in \dot{C}^1$ we define

$$(2.1.2) \quad \|v\|_{Y_\lambda} = \left\| -\frac{\dot{v}}{\lambda} + \lambda v \right\|_X,$$

where we denoted $\dot{v} = \frac{dv}{dt}$. Let Y_λ be the completion of \dot{C}^1 with respect to the norm $\|\cdot\|_{Y_\lambda}$.

Remark 2.1.1. In [1] we have introduced in \dot{C}^1 the norm

$$(2.1.3) \quad \|v\|_{Z_\lambda} = \left[\|\dot{v}\|_X^2 + \|\lambda v\|_X^2 \right]^{1/2}$$

and have shown (Lemma 2.1 of [1]) that

$$(2.1.4) \quad C_1 \|v\|_{Z_\lambda} \leq \|v\|_{Y_\lambda} \leq C_2 \|v\|_{Z_\lambda}$$

with $C_1 > 0$, $C_2 < +\infty$ independent of λ and v but dependent on T .

On $X \times Y_\lambda$ we define the bilinear form

$$(2.1.3) \quad B_\lambda(u, v) = \int_0^T u \left(-\frac{\dot{v}}{\lambda} + \lambda v \right) dt.$$

Further, let $F \in Y'_\lambda$ be a linear functional on Y_λ . We can define

Problem P_λ . For given $F \in Y'_\lambda$ find $u_0 \in X$ such that

$$(2.1.4) \quad B_\lambda(u_0, v) = F(v) \quad \forall v \in Y_\lambda.$$

This problem has been analyzed in [1] where, among others, the unique solvability of it has been proved. Moreover, the solution of the problem P_λ is a weak solution u_0 of the initial value problem

$$(2.1.5) \quad \begin{aligned} \frac{\dot{u}}{\lambda} + \lambda u &= f \\ u(0) &= a\lambda, \end{aligned}$$

if

$$F(v) = \int_0^T f v dt + a v(0).$$

Let now $k \geq 0$, an integer, $A = (t_1, t_2)$

$$S^k(A) = \{w \mid w \text{ is polynomial of degree } \leq k \text{ on } A\}$$

$$S^k(A) = \{w \in S^k(A) \mid w(t_2) = 0\}.$$

We will first consider the following auxiliary problems:

for any $\lambda > 0$, $q \geq 1$, $k = 0, 1, \dots$, $c > 0$, find $\omega^{[k]} \in S^k(-1, 1)$ such that

$$(2.1.6) \quad -\frac{\dot{\omega}^{[k]}}{\lambda} + \lambda \omega^{[k]} = c P_k,$$

where P_k is the Legendre polynomial of degree k .

The following lemmas give us the estimates of the solutions for the above problems which are important in our further analysis.

Lemma 2.1.1. The problems (2.1.6) have the unique solutions $\omega^{[k]}$ for any $k = 0, 1, \dots$. They satisfy:

- i) $\| -\frac{\dot{\omega}^{[k]}}{\lambda} + \lambda \omega^{[k]} \|_{L_2(-1, 1)} = c \sqrt{\frac{2}{2k+1}}$
- ii) $\omega^{[k]}(1) > 0$ and $|\omega^{[k]}(-1)| \leq \omega^{[k]}(1)$
- iii) if $k_1 \leq k_2$ then $\omega^{[k_1]}(1) \leq \omega^{[k_2]}(1)$.

Proof. Let us represent the solution $\omega^{[k]}$ of the problem (2.1.6) by

$$(2.1.7) \quad \omega^{[k]}(t) = \sum_{j=0}^k \beta_j^{[k]} P_j(t),$$

where P_j is the Legendre polynomial of degree j and use the summation formula (see, e.g., [2])

$$(2.1.8) \quad \dot{P}_j(t) = \sum_{i=0}^j (2j-4i-1) P_{j-2i-1}(t).$$

Then we get the system of linear equations with the nonsingular matrix

$$(2.1.9) \quad \begin{bmatrix} \lambda & 0 & 0 & 0 & \dots & 0 \\ -\frac{(2k-1)}{\lambda} & \lambda & 0 & 0 & \dots & 0 \\ 0 & -\frac{(2k-3)}{\lambda} & \lambda & 0 & \dots & 0 \\ -\frac{(2k-5)}{\lambda} & 0 & -\frac{(2k-5)}{\lambda} & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \dots & 0 & -\frac{1}{\lambda} & 0 & -\frac{1}{\lambda} & \lambda \end{bmatrix} \begin{bmatrix} \beta_k^{[k]} \\ \beta_{k-1}^{[k]} \\ \beta_{k-2}^{[k]} \\ \beta_{k-3}^{[k]} \\ \vdots \\ \beta_0^{[k]} \end{bmatrix} = \begin{bmatrix} c \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

So, we can see that the representation (2.1.7) is unique and the problem (2.1.6) has a unique solution for any $k = 0, 1, 2, \dots$. Then

$$i) \quad \left\| -\frac{\dot{\omega}^{[k]}}{\lambda} + \lambda \omega^{[k]} \right\|_{L_2(-1,1)} = \|cP_k\|_{L_2(-1,1)} = c\sqrt{\frac{2}{2k+1}}.$$

Furthermore, all the coefficients $\beta_j^{[k]}$, $j = 0, 1, \dots, k$, are positive, which immediately implies (ii). Indeed, from the property of the Legendre polynomials, we have

$$|\omega^{[k]}(-1)| = \left| \sum_{j=0}^k (-1)^j \beta_j^{[k]} \right| \leq \sum_{j=0}^k \beta_j^{[k]} = \omega^{[k]}(1).$$

By analyzing the system (2.1.9) for k and $k+1$ we can observe that

$$\beta_{k+1}^{[k+1]} = \beta_k^{[k]} = \frac{1}{\lambda},$$

$$\beta_{k+1-j}^{[k+1]} > \beta_{k-j}^{[k]} \quad \text{for } j = 1, 2, \dots, k,$$

and

$$\beta_0^{[k+1]} > 0.$$

The above inequalities imply that

$$\sum_{k=0}^{k_2} \beta_j^{[k_2]} > \sum_{k=0}^{k_1} \beta_j^{[k_1]} \quad \text{if } k_2 > k_1,$$

i.e.,

$$\omega^{[k_1]}(1) < \omega^{[k_2]}(1) \quad \text{if } k_1 < k_2.$$

Lemma 2.1.2. Let $\omega_1^{[k_1]}$ and $\omega_2^{[k_2]}$ be the solutions of

$$-\frac{\dot{\omega}_1^{[k_1]}}{c_1 \lambda} + c_1 \lambda \omega_1^{[k_1]} = c_1 P_{k_1}, \quad t \in (-1, 1)$$

and

$$-\frac{\dot{\omega}_2^{[k_2]}}{c_2 \lambda} + c_2 \lambda \omega_2^{[k_2]} = c_2 P_{k_2}, \quad t \in (-1, 1),$$

respectively. Then, for $k_1 \leq k_2$ and $c_2/c_1 = \sigma$,

$$(2.1.10) \quad \omega_1^{[k_1]}(1) \leq \max(1, \sigma^{2k_2}) \omega_2^{[k_2]}(1).$$

Proof. The values $\omega_1^{[k_1]}(1)$ and $\omega_2^{[k_2]}(1)$ can be expressed in the following way (see (2.1.9)):

$$\omega_1^{[k_1]}(1) = \frac{1}{\lambda} \left[1 + \sum_{i=1}^{k_1} \frac{f(i, k_1)}{(c_1 \lambda)^{2i}} \right],$$

$$\omega_2^{[k_2]}(1) = \frac{1}{\lambda} \left[1 + \sum_{i=1}^{k_2} \frac{f(i, k_2)}{(c_2 \lambda)^{2i}} \right],$$

with some integer function $f(i, k) \geq 1$. Then

$$\begin{aligned} \omega_2^{[k_2]}(1) &= \frac{1}{\lambda} \left[1 + \sum_{i=1}^{k_2} \frac{f(i, k_2)}{(c_2 \lambda)^{2i}} \right] = \frac{1}{\lambda} \left[1 + \sum_{i=1}^{k_2} \frac{f(i, k_2)}{(\sigma c_1 \lambda)^{2i}} \right] \\ &\geq \frac{1}{\max(1, \sigma^{2k_2}) \lambda} \left[1 + \sum_{i=1}^{k_2} \frac{f(i, k_2)}{(c_1 \lambda)^{2i}} \right]. \end{aligned}$$

Now, by using Lemma 2.1.1 with $\lambda \rightarrow \lambda c_1$ and $c = c_1$ we complete the proof. ■

2.2. The h-p version of the finite element method for the problem P_λ

Let now

$$(2.2.1) \quad \Delta_N : 0 = t_0 < t_1 < t_2 < \dots < t_N = T$$

be a subdivision of $(0, T)$ into time intervals $I_n = (t_{n-1}, t_n)$, $n = 1, 2, \dots, N$, of the length $\tau_n = t_n - t_{n-1}$, and $d_n = \tau_n/2$. Furthermore, let $\underline{q} = (q_1, q_2, \dots, q_N)$, $q_n \geq 1$, an integer, $n = 1, 2, \dots, N$, and

$$(2.2.3) \quad S = S^{\underline{q}-1} = \{u \in X \mid u|_{I_n} \in S^{q_n-1}(I_n)\}$$

$$(2.2.4) \quad V = S^{\underline{q}} = \{v \in Y_\lambda \mid v|_{I_n} \in S^{q_n}(I_n)\}.$$

The definition of Y_λ implies the continuity of $v \in V$ and $v(T) = 0$.

We define the h-p version for the problem P_λ : For given $F \in Y'_\lambda$, a mesh Δ_N , $\underline{q} = (q_1, q_2, \dots, q_N)$ find $u_{\underline{q}} \in S$ such that

$$(2.2.5) \quad B_\lambda(u_{\underline{q}}, v) = F(v), \quad \forall v \in V.$$

Remark 2.2.1. The problem (2.2.5) is actually the set of N problems to be solved in the succession. Let us denote for $n = 1, 2, \dots, N$

$$S_n = \{u_n : I_n \rightarrow \mathbb{R} \mid u_n \text{ is a polynomial of degree } \leq q_n - 1\},$$

$$V_n = \{v_n : I_n \rightarrow \mathbb{R} \mid v_n \text{ is a polynomial of degree } \leq q_n\}, \quad n = 1, \dots, N-1,$$

$$V_N = \{v_N : I_N \rightarrow \mathbb{R} \mid v_N \text{ is a polynomial of degree } \leq q_N, v_N(t_N) = 0\}, \quad n = N,$$

and represent the space V_n in the form of the direct sum

$$(2.2.6) \quad V_n = V_n^{[1]} \oplus V_n^{[2]}, \quad n = 1, 2, \dots, N-1, \quad V_N = V_N^{[1]}$$

where

$$V_n^{[1]} = \{v_n^{[1]} \in V_n \mid v_n^{[1]}(t_n) = 0\}$$

$$V_n^{[2]} = \{v_n^{[2]} \in V_n \mid v_n^{[2]} = c_n \chi_n\},$$

where χ_n is a (fixed but arbitrary) polynomial of degree $\leq q_n$ such that $\chi_n(t_{n-1}) = 0$, $\chi_n(t_n) = 1$.

Taking into account the continuity of each function $v \in V$, $v|_{I_n} \in V_n$ we see that

$$\dim V = \dim S = \sum_{n=1}^N q_n.$$

The result is that the problem (2.2.5) is replaced by the following set of N problems

$$(2.2.7) \quad B_{\lambda}^n(u_n, v_n^{[1]}) = F^n(v_n^{[1]}) \quad \forall v_n^{[1]} \in V_n^{[1]},$$

where for $n = 1, 2, \dots, N$,

$$B_{\lambda}^n(u_n, v_n^{[1]}) = \int_{I_n} u_n \left[-\frac{\dot{v}_n^{[1]}}{\lambda} + \lambda v_n^{[1]} \right] dt$$

with

$$(2.2.9) \quad F^1(v_1^{[1]}) = \int_{I_1} f v_1^{[1]} dt + a v_1^{[1]}(0)$$

$$F^n(v_n^{[1]}) = \int_{I_n} f v_n^{[1]} dt + \int_{I_{n-1}} f v_{n-1}^{[2]} dt - B_{\lambda}^{n-1}(u_{n-1}, v_{n-1}^{[2]}), \quad n = 2, 3, \dots, N,$$

and

$$v_n^{[2]}(t_n) = v_{n+1}^{[1]}(t_n) \quad \forall n = 1, 2, \dots, N-1.$$

Let us note that the above sequence of problems is independent of the selection of χ_n .

Theorem 2.2.1. 1) Let $u \in S$, $v \in V$, then

$$(2.2.10) \quad |B_\lambda(u, v)| \leq \|u\|_X \|v\|_{Y_\lambda}$$

ii) If $q_1 \geq q_2 \geq \dots \geq q_N \geq 1$ and $\tau_1/\tau_j \leq \sigma$ for any $1 \leq j \leq N$, then

$$(2.2.11) \quad d_\lambda(q, N) = \inf_{\substack{u \in S \\ \|u\|_X=1}} \sup_{\substack{v \in V \\ \|v\|_{Y_\lambda} \leq 1}} |B_\lambda(u, v)| \geq \frac{1}{4} q_1^{-1/2} N^{-1} \min(1, \sigma^{-(2q_1+(1/2))})$$

iii) Let $v \in V$, $v \neq 0$, then

$$(2.2.12) \quad \sup_{\substack{u \in S \\ \|u\|_X=1}} |B_\lambda(u, v)| > 0.$$

Proof. i) (2.2.10) follows from the Schwarz inequality.

ii) Denoting by

$$(2.2.13) \quad \|u_n\|_{X,n} = \left(\int_{I_n} u_n^2 dt \right)^{1/2}$$

and

$$(2.2.14) \quad \|v_n\|_{Y_\lambda,n} = \left(\int_{I_n} \left(-\frac{\dot{v}_n}{\lambda} + \lambda v_n \right)^2 dt \right)^{1/2}$$

we get

$$(2.2.15) \quad \|u\|_X = \left(\sum_{n=1}^N \|u_n\|_{X,n}^2 \right)^{1/2}$$

and

$$(2.2.16) \quad \|v\|_{Y_\lambda} = \left(\sum_{n=1}^N \|v_n\|_{Y_\lambda,n}^2 \right)^{1/2}.$$

For given $u \in S$, $u|_{I_n} = u_n \in S_n$, we will construct $\bar{v} \in V$, such that

$$\bar{v}|_{I_n} = \bar{v}_n \in V_n.$$

and using $v = \bar{v}$ as the function from the test space we will prove our claim.

Let

$$(2.2.17) \quad \omega_n = \sum_{k=1}^{q_n-1} a_{k,n}^n \omega_n^{[k]},$$

where

$$(2.2.18) \quad u_n = \sum_{k=1}^{q_n-1} a_{k,n}^n P_k^n,$$

where P_k^n is the k -th Legendre polynomial on I_n and $\omega_n^{[k]} \in S^k(I_n)$ is the unique solution of

$$(2.2.19) \quad -\frac{\dot{\omega}_n^{[k]}}{\lambda} + \lambda \omega_n^{[k]} = P_k^n, \quad t \in I_n, \quad k = 0, 1, \dots$$

By transforming these problems onto the reference element $(-1, 1)$ we can see that

$$\tilde{\omega}_n^{[k]} \in S^k(-1, 1), \quad k = 0, 1, \dots, q_n - 1$$

satisfy the following equations

$$(2.2.20) \quad -\frac{\dot{\tilde{\omega}}_n^{[k]}}{\lambda d_n^{1/2}} + \lambda d_n^{1/2} \tilde{\omega}_n^{[k]} = d_n^{1/2} P_k, \quad t \in (-1, 1), \quad k = 0, 1, \dots, q_n - 1.$$

The problems (2.2.20) are of the type (2.1.6) with different $\lambda_n = \lambda d_n^{1/2}$ and $c_n = d_n^{1/2}$ for different $n = 1, 2, \dots, N$. Thus

$$(2.2.21) \quad B_{\lambda}^n(u_n, \omega_n) = \int_{I_n} u_n \left(-\frac{\dot{\omega}_n}{\lambda} + \lambda \omega_n \right) dt = \|u_n\|_{X,n}^2$$

and

$$(2.2.22) \quad \|\omega_n\|_{Y_{\lambda},n} = \|u_n\|_{X,n}.$$

Now let

$$(2.2.23) \quad w_n = \omega_n - \frac{\omega_n(t_n)}{\omega_n^{[q_n]}(t_n)} \omega_n^{[q_n]},$$

where $\omega_n^{[q_n]}$ is the solution of (2.2.19) with $k = q_n$. Obviously,

$$(2.2.24) \quad w_n(t_n) = 0, \quad n = 1, 2, \dots, N,$$

and

$$B_\lambda^n(u_n, w_n) = B_\lambda^n\left(u_n, \omega_n - \frac{\omega_n(t_n)}{\omega_n^{[q_n]}(t_n)} \omega_n^{[q_n]}\right) = B_\lambda^n(u_n, \omega_n) = \|u_n\|_{X,n}^2$$

since

$$B_\lambda^n(u_n, \omega_n^{[q_n]}) = 0.$$

Further

$$(2.2.26) \quad |\omega_n(t_n)| = \left| \sum_{k=0}^{q_n-1} a_k^n \omega_n^{[k]}(t_n) \right| = \left| \sum_{k=0}^{q_n-1} \sqrt{\frac{2}{2k+1}} a_k^n \sqrt{\frac{2k+1}{2}} \omega_n^{[k]}(t_n) \right|$$

$$\leq \left[\sum_{k=0}^{q_n-1} \frac{2}{2k+1} (a_k^n)^2 \right]^{1/2} \left[\sum_{k=0}^{q_n-1} \frac{2k+1}{2} (\omega_n^{[k]}(t_n))^2 \right]^{1/2}$$

$$\leq d_n^{-1/2} 2^{-1/2} q_n |\omega_n^{[q_n]}(t_n)| \|u_n\|_{X,n},$$

where we used the Schwarz inequality and Lemma 2.1.1 iii). Thus

$$(2.2.27) \quad \|w_n\|_{Y_\lambda, n} \leq \|\omega_n\|_{Y_\lambda, n} + \frac{|\omega_n(t_n)|}{|\omega_n^{[q_n]}(t_n)|} \|\omega_n^{[q_n]}\|_{Y_\lambda, n}$$

$$\leq \|u_n\|_{X,n} + \frac{d_n^{-1/2} q_n |\omega_n^{[q_n]}(t_n)| \|u_n\|_{X,n} d_n^{1/2}}{\sqrt{2} |\omega_n^{[q_n]}(t_n)|} \sqrt{\frac{2}{2q_n+1}}$$

$$\leq \|u_n\|_{X,n} \left(1 + \frac{q_n}{\sqrt{2q_n+1}} \right) \leq 2q_n^{1/2} \|u_n\|_{X,n}.$$

Since $|\omega_n^{[k]}(t_{n-1})| < \omega_n^{[k]}(t_n)$ for $k = 0, 1, \dots, q_n$ (Lemma 2.1.1 ii), we get analogously as before

$$(2.2.28) \quad |w_n(t_{n-1})| \leq |\omega_n(t_{n-1})| + \frac{|\omega_n(t_n)|}{|\omega_n^{[q_n]}(t_n)|} |\omega_n^{[q_n]}(t_{n-1})| \\ \leq d_n^{-1/2} \sqrt{2} q_n |\omega_n^{[q_n]}(t_n)| \|u_n\|_{X,n}.$$

Let us now define the following sequence of functions $z_n \in V_n$. We put

$$(2.2.29) \quad z_N = 0$$

and

$$z_{N-1} \in V_{N-1}$$

be the solution of the problem

$$-\frac{\dot{z}_{N-1}}{\lambda} + \lambda z_{N-1} = \frac{w_N(t_{N-1})}{\omega_{N-1}^{[q_{N-1}]}(t_{N-1})} p_{N-1}^{N-1}, \quad t \in I_{N-1},$$

where $\omega_{N-1}^{[q_{N-1}]}$ is defined by (2.2.19). Then

$$(2.2.31) \quad \|z_{N-1}\|_{Y_{\lambda,N-1}} \leq \frac{|w_N(t_{N-1})|}{|\omega_{N-1}^{[q_{N-1}]}(t_{N-1})|} \|p_{N-1}^{N-1}\|_{X,N-1} \\ \leq \frac{d_N^{-1/2} \sqrt{2} q_N |\omega_N^{[q_N]}(t_N)| \|u_N\|_{X,N} d_{N-1}^{1/2} \sqrt{2}}{|\omega_{N-1}^{[q_{N-1}]}(t_N)| \sqrt{2} q_{N-1} + 1}.$$

If $q_N \leq q_{N-1}$ and $d_{N-1}/d_N = \sigma_{N,N-1} \leq \sigma$, then from Lemma 2.1.2

$$(2.2.32) \quad |\omega_N^{[q_N]}(t_N)| \leq |\omega_{N-1}^{[q_{N-1}]}(t_N)| \max(1, \sigma^{2q_{N-1}})$$

and

$$(2.2.33) \quad \|z_{N-1}\|_{Y_{\lambda,N-1}} \leq 2^{1/2} q_{N-1}^{1/2} \max(1, \sigma^{2q_{N-1}+1/2}) \|u_N\|_{X,N}.$$

Further,

$$z_{N-2} \in V_{N-2}$$

is the solution of the problem

$$(2.2.34) \quad -\frac{\dot{z}_{N-2}}{\lambda} + \lambda z_{N-2} = \frac{w_{N-1}(t_{N-2}) + z_{N-1}(t_{N-2})}{\omega_{N-2}^{[q_{N-2}]}(t_{N-2})} p_{q_{N-2}}^{N-2}, \quad t \in I_{N-2}.$$

Since

$$|z_{N-1}(t_{N-2})| < |z_{N-1}(t_{N-1})| = |w_N(t_{N-1})|$$

and from (2.2.28)

$$(2.2.35) \quad |w_{N-1}(t_{N-2})| \leq d_{N-1}^{-1/2} \sqrt{2} q_{N-1} |\omega_{N-1}^{[q_{N-1}]}(t_{N-1})| \|u_{N-1}\|_{X, N-1}$$

while

$$(2.2.36) \quad |w_N(t_{N-1})| \leq d_N^{-1/2} \sqrt{2} q_N |\omega_N^{[q_N]}(t_N)| \|u_N\|_{X, N},$$

we have (with Lemma 2.1.2)

$$(2.2.37) \quad \|z_{N-2}\|_{Y_{\lambda, N-2}} \leq \left[\frac{d_{N-1}^{-1/2} \sqrt{2} q_{N-1} |\omega_{N-1}^{[q_{N-1}]}(t_{N-1})| \|u_{N-1}\|_{X, N-1}}{|\omega_{N-2}^{[q_{N-2}]}(t_{N-2})|} + \frac{d_N^{-1/2} \sqrt{2} q_N |\omega_N^{[q_N]}(t_N)| \|u_N\|_{X, N}}{|\omega_{N-2}^{[q_{N-2}]}(t_{N-2})|} \right] d_{N-2}^{1/2} \sqrt{\frac{2}{2q_{N-2}+1}} \\ \leq 2^{1/2} q_{N-2}^{1/2} \max(1, \sigma^{2q_{N-2}+(1/2)}) (\|u_{N-1}\|_{X, N-1} + \|u_N\|_{X, N}).$$

Hence we can define the sequence $z_n \in V_n$ in the recursive way

$$(2.2.38) \quad z_N = 0, \quad t \in I_N,$$

$$-\frac{\dot{z}_{n-1}}{\lambda} + \lambda z_{n-1} = \frac{w_n(t_{n-1}) + z_n(t_{n-1})}{\omega_{n-1}^{[q_{n-1}]}(t_{n-1})} p_{q_{n-2}}^{n-2}, \quad t \in I_{n-1}, \quad n = N, N-1, N-2, \dots, 2,$$

where $\omega_{n-1}^{[q_{n-1}]}$ is defined by (2.2.19). Then

$$(2.2.39) \quad \|z_{n-1}\|_{Y_{\lambda}, n-1} \leq 2^{1/2} q_{n-1}^{1/2} \max(1, \sigma^{2q_{n-1}+(1/2)}) \sum_{i=n}^N \|u_i\|_{X, i}.$$

Finally, let $\bar{v} \in V$ such that

$$(2.2.40) \quad \bar{v}|_{I_n} = w_n + z_n \quad \text{for } n = 1, 2, \dots, N.$$

Obviously from our construction, \bar{v} is continuous and $\bar{v}(T) = 0$. Moreover,

$$B_{\lambda}^n(u_n, z_n) = 0 \quad \text{for all } n = 1, 2, \dots, N.$$

Thus

$$(2.2.41) \quad \sup_{\substack{v \in V \\ \|v\|_{Y_{\lambda}} \leq 1}} |B_{\lambda}(u, v)| \geq \frac{\|u_1\|_{X, 1}^2 + \|u_2\|_{X, 2}^2 + \dots + \|u_N\|_{X, N}^2}{(\|w_1 + z_1\|_{Y_{\lambda}, 1}^2 + \dots + \|w_N + z_N\|_{Y_{\lambda}, N}^2)^{1/2}} \\ \geq \frac{\|u_1\|_{X, 1}^2 + \|u_2\|_{X, 2}^2 + \dots + \|u_N\|_{X, N}^2}{2^{1/2} \left[\|w_1\|_{Y_{\lambda}, 1}^2 + \dots + \|w_N\|_{Y_{\lambda}, N}^2 + \|z_1\|_{Y_{\lambda}, 1}^2 + \dots + \|z_{N-1}\|_{Y_{\lambda}, N-1}^2 \right]^{1/2}} \\ \geq \frac{\|u_1\|_{X, 1}^2 + \dots + \|u_N\|_{X, N}^2}{2^{3/2} q_1^{1/2} \left[\|u_1\|_{X, 1}^2 + \dots + \|u_N\|_{X, N}^2 + \max(1, \sigma^{4q_1+1}) \sum_{n=1}^N \left[\sum_{i=n}^N \|u_i\|_{X, i}^2 \right] \right]^{1/2}}$$

and further

$$(2.2.42) \quad \sup_{\substack{v \in V \\ \|v\|_{Y_{\lambda}} = 1}} |B_{\lambda}(u, v)| \geq \frac{\sum_{n=1}^N \|u_n\|_{X, n}^2}{2^{3/2} q_1^{1/2} \left[\sum_{n=1}^N \|u_n\|_{X, n}^2 + N(N-1) \max(1, \sigma^{4q_1+1}) \sum_{n=1}^N \|u_n\|_{X, n}^2 \right]^{1/2}} \\ \geq \frac{\left[\sum_{n=1}^N \|u_n\|_{X, n}^2 \right]^{1/2}}{4q_1^{1/2} N \max(1, \sigma^{2q_1+1/2})}$$

The inequality (2.2.42) immediately yields ii).

iii) For any $v \in V$, $v_n = v|_{I_n} \in V_n$, we select $\bar{u} \in S$, such that

$$(2.2.43) \quad \bar{u}_n = \bar{u}|_{I_n} = -\dot{v}_n.$$

Then from continuity of $v \in V$ and $v(T) = 0$,

$$\begin{aligned} (2.2.44) \quad B(\bar{u}, v) &= \int_0^T \bar{u} \left(-\frac{\dot{v}}{\lambda} + \lambda v \right) dt = \sum_{n=1}^N \int_{I_n} (-\dot{v}_n) \left(-\frac{\dot{v}_n}{\lambda} + \lambda v_n \right) dt \\ &= \sum_{n=1}^N \left[\int_{I_n} \frac{\dot{v}_n^2}{\lambda} dt - \lambda \int_{I_n} \dot{v}_n v_n dt \right] = \sum_{n=1}^N \int_{I_n} \frac{\dot{v}_n^2}{\lambda} dt - \frac{\lambda}{2} \sum_{n=1}^N \int_{I_n} \frac{d}{dt} (v_n^2) dt \\ &= \sum_{n=1}^N \int_{I_n} \frac{\dot{v}_n^2}{\lambda} dt - \frac{\lambda}{2} \sum_{n=1}^N [v_n^2(t_n) - v_n^2(t_{n-1})] \\ &= \sum_{n=1}^N \int_{I_n} \frac{\dot{v}_n^2}{\lambda} dt - \frac{\lambda}{2} [v_N^2(t_N) - v_1^2(t_0)] \\ &= \sum_{n=1}^N \int_{I_n} \frac{\dot{v}_n^2}{\lambda} dt + \frac{\lambda}{2} v_1^2(t_0) > 0 \quad \text{if } v \neq 0. \quad \blacksquare \end{aligned}$$

Theorem 2.2.1 together with Theorem 2.2 of [1] and Theorem 6.2.1 of [3] (see also Appendix of [1]) yields

Theorem 2.2.2. There is a unique $u_{\underline{q}}$ satisfying (2.2.5). If $u_0 \in X$ is the exact solution of the problem P_{λ} and $q_1 \geq q_2 \geq \dots \geq q_N \geq 1$, $\tau_1/\tau_j \leq \sigma$ for $1 \leq i < j \leq N$, then

$$(2.2.45) \quad \|u_0 - u_{\underline{q}}\|_X \leq \left(1 + 4q_1^{1/2} N \max(1, \sigma^{2q_1 + (1/2)}) \right) \inf_{w \in S} \|u_0 - w\|_X.$$

2.3. Comments.

If (2.2.11) is optimal, i.e., if there is a sequence of $\{\lambda_1, \underline{q}_1, N_1\}$ such that

$$d_{\lambda_1}(q_1, N_1) \leq C q_{1,1}^{-1/2} N_1^{-1} \min(1, \sigma^{-(2q_{1,1} + (1/2))})$$

with C independent of i , then there is a sequence of solutions $u_i \in X$ such that

$$\frac{\|u_i - u_{q_i}\|_X}{\inf_{w \in S} \|u_i - w\|_X} \geq C q_{1,1}^{1/2} N_1^{\max(1, \sigma^{2q_{1,1} + (1/2)})}.$$

(See Theorem 2.10 of [4].)

The optimality with respect to the exponent of q has been proved in [1]. Let us now prove the optimality with respect to the exponent of N . Consider the case of a uniform mesh

$$(2.3.1) \quad \Delta : 0 = t_0 < t_1 < \dots < t_N = T$$

with

$$(2.3.2) \quad t_n - t_{n-1} = \tau = \frac{1}{N}.$$

Let $q_n = q = 1$ and $\bar{v} \in V$ be defined by (Figure 2.3.1)

$$(2.2.3) \quad \bar{v}(t) = (-1)^n \left[(2N - 2n + 1)t + \frac{N - n - (2N - 2n + 1)n}{N} \right], \quad t \in I_n, \quad n = 1, 2, \dots, N.$$

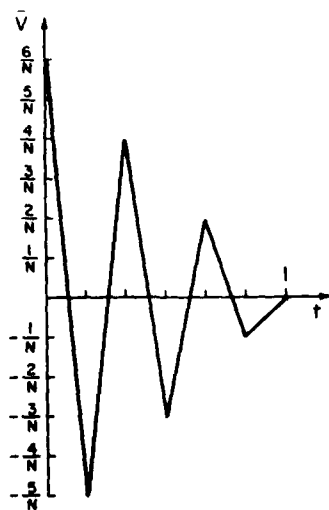


Figure 2.3.1 $\bar{v}(t)$ for $N = 6$.

Now, let $\psi_1, \psi_2, \dots, \psi_{N-1} \in S$ denote the basis function defined by

$$(2.3.4) \quad \psi_n(t) = \begin{cases} 1 & t \in I_n \\ 0 & t \notin I_n \end{cases} \quad n = 1, 2, \dots, N.$$

Then we have

$$(2.3.5) \quad B_\lambda(\psi_n, \bar{v}) = \int_{I_n} \psi_n \left[-\frac{\dot{\bar{v}}}{\lambda} + \lambda \bar{v} \right] dt = (-1)^n \left[\frac{2N-2n+1}{N\lambda} + \frac{\lambda}{2N^2} \right], \quad n = 1, 2, \dots, N.$$

Any $u \in S$ can be written in the form

$$(2.3.6) \quad u = \sum_{n=1}^N c_n \psi_n \quad \text{with} \quad \|u\|_X^2 = \frac{1}{N} \sum_{n=1}^N c_n^2,$$

then

$$(2.3.7) \quad B_\lambda(u, \bar{v}) = \sum_{n=1}^N \left[(-1)^{n+1} c_n \left(\frac{2N-2n+1}{N\lambda} + \frac{\lambda}{2N^2} \right) \right].$$

Obviously

$$(2.3.8) \quad \sup_{\|u\|_X \leq 1} \left| \sum_{n=1}^N \left[(-1)^{n+1} c_n \left(\frac{2N-2n+1}{N\lambda} + \frac{\lambda}{2N^2} \right) \right] \right| = \left[N \sum_{n=1}^N \left[\frac{2N-2n+1}{N\lambda} + \frac{\lambda}{2N^2} \right]^2 \right]^{1/2} \\ \leq K_1 \left(\frac{N^2}{\lambda^2} + 1 + \frac{\lambda^2}{N^2} \right)^{1/2}.$$

Note that there exists a constant $K_2 > 0$ such that

$$(2.3.9) \quad \|\bar{v}\|_{Y_\lambda} \geq K_2 \left(\frac{N^2}{\lambda^2} + 1 + \lambda^2 \right)^{1/2}.$$

Combining (2.3.8) and (2.3.9) we see that if λ is sufficiently large in comparison with N (e.g., $\lambda \geq N$), there is a constant $C > 0$ such that

$$(2.3.10) \quad \inf_{\substack{v \in V \\ \|v\|_{Y_\lambda} = 1}} \sup_{\substack{u \in S \\ \|u\|_X \leq 1}} |B_\lambda(u, v)| \leq CN^{-1}.$$

We have to understand (2.3.10) as an estimate which is uniform with respect to λ . We have seen that for given N , the estimate (2.3.10) holds for λ sufficiently large respectively to N , e.g., $\lambda \geq N$.

On the other hand we can easily show that

$$(2.3.11) \quad \inf_{\substack{v \in V \\ \|v\|_{Y_\lambda} = 1}} \sup_{\substack{u \in S \\ \|u\|_X \leq 1}} |B_\lambda(u, v)| \geq C(1+\lambda^4)^{-1}.$$

In fact, given $v \in V$, select $u = -\frac{\dot{v}}{\lambda}$. Then $u \in S$ and we have

$$(2.3.12) \quad B_\lambda(u, v) = \int_0^T \left(\frac{v^2}{\lambda^2} - \dot{v}v \right) dt = \int_0^T \frac{v^2}{\lambda^2} dt + \frac{1}{2}v^2(0) \geq \int_0^T \frac{v^2}{\lambda^2} dt.$$

Thus, using Remark 2.1.1

$$(2.3.13) \quad \|u\|_X \leq c\|v\|_{Y_\lambda}$$

and

$$(2.3.14) \quad \int_0^T v^2 dt \leq C \int_0^T \dot{v}^2 dt \quad (\text{because } v(T) = 0),$$

$$(2.3.15) \quad \|v\|_{Y_\lambda}^2 \leq C \left[\int_0^T \left(\frac{\dot{v}^2}{\lambda^2} + \lambda^2 v^2 \right) dt \right] \leq C(1+\lambda^4) \int_0^T \frac{\dot{v}^2}{\lambda^2} dt.$$

Hence

$$(2.3.16) \quad B_\lambda(u, v) \geq \frac{C}{1+\lambda^4} \|v\|_{Y_\lambda}^2$$

which yields (2.3.11).

The assumption $q_1 \geq q_2 \geq \dots \geq q_N$ seems to be only implied by the

specific construction of the sequence $\{z_n\}_{n=1}^N$ in the proof of Theorem 2.2.1.

However, the counterexample considered below will show that the assumption $\tau_1/\tau_j \leq \sigma$ for $1 \leq j \leq N$ is necessary, i.e., without this condition $d_\lambda(q, N)$ can be arbitrary small (for some λ). Let us consider only two elements of length 1 and τ , respectively, and $q_1 = q_2 = 1$.

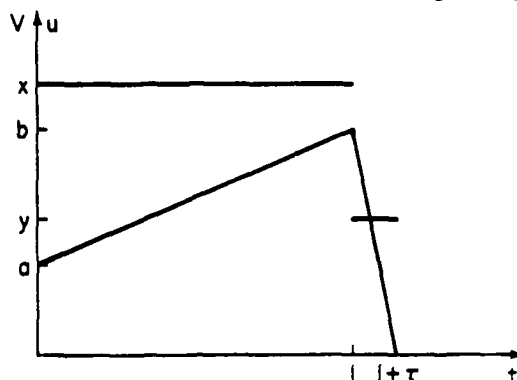


Figure 2.3.2

Thus

$$(2.3.17) \quad v(t) = \begin{cases} (b-a)t + a & 0 < t \leq 1 \\ -\frac{b}{\tau}t + b(1 + \frac{1}{\tau}) & 1 < t < 1+\tau \end{cases}$$

$$(2.3.18) \quad u(t) = \begin{cases} x & 0 < t < 1 \\ y & 1 < t < 1+\tau. \end{cases}$$

Then

$$(2.3.19) \quad B(u, v) = \frac{x(a-b)}{\lambda} + \frac{yb}{\lambda} + \lambda \left[\frac{x(a+b)}{2} + \frac{yb\tau}{2} \right]$$

and

$$(2.3.20) \quad \|u\|_X^2 = x^2 + y^2\tau$$

$$\|v\|_{Z_\lambda}^2 = \int_0^{1+\tau} \left[\left(\frac{\dot{v}}{\lambda} \right)^2 + (\lambda v)^2 \right] dt = \frac{a^2 - ab + b^2}{\lambda^2} + \frac{b^2}{\tau\lambda^2} + \lambda^2 \left(\frac{a^2 - ab + b^2}{3} + \frac{b^2\tau}{3} \right).$$

Obviously,

$$\begin{aligned}
 (2.3.21) \quad R &= \lambda \sup_{\substack{[x,y] \\ x^2+y^2=1}} |x \frac{a+b}{2} + y \sqrt{\tau} \frac{b\sqrt{\tau}}{2}| = \lambda \left(\frac{(a+b)^2}{4} + \frac{b^2 \tau}{4} \right)^{1/2} \\
 &= \lambda \left(\frac{1}{4} a^2 + \frac{1}{4} 2ab + b^2 \frac{\tau+1}{4} \right)^{1/2}.
 \end{aligned}$$

Let

$$(2.3.22) \quad \inf \lambda^2 \left(\frac{1}{4} a^2 + \frac{1}{4} 2ab + b^2 \frac{\tau+1}{4} \right) = Q_{\min},$$

where \inf is taken over a, b such that

$$(2.3.23) \quad \lambda^2 \left(\frac{1}{3} a^2 + \frac{1}{6} 2ab + b^2 \frac{\tau+1}{3} \right) = 1.$$

Then Q_{\min} is the smallest eigenvalue of the following eigenvalue problem

$$(2.3.24) \quad AZ = QBZ,$$

where $Z = (a, b)^T$,

$$(2.3.25) \quad A = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1+\tau}{4} \end{bmatrix} \quad B = \begin{bmatrix} \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1+\tau}{3} \end{bmatrix},$$

$$\det(A - QB) = Q^2 \frac{4\tau+3}{36} - Q \frac{2\tau+1}{12} + \frac{\tau}{16} = 0$$

and $Q_{\min} \rightarrow 0$ as $\tau \rightarrow 0$. Assuming that (2.3.23) holds, it is obvious that there exists a function $0 < \mathcal{H}(\tau) < \infty$, $0 < \tau < \infty$, such that

$$(2.3.26) \quad \|v\|_{Z_\lambda}^2 \leq \mathcal{H}(\tau) \frac{1}{\lambda^4} + 1$$

and

$$(2.3.27) \quad \sup_{\|u\|_X \leq 1} |B_\lambda(u, v)| \leq \mathcal{H}(\tau) \frac{1}{\lambda^3} + R.$$

Hence, using Remark 2.1.1, we get

$$d_{\lambda}(\tau) = \inf_{\|v\|_{Y_{\lambda}}=1} \sup_{\|u\|_X \leq 1} |B_{\lambda}(u, v)| \leq \left[\frac{\mathcal{H}(\tau)}{\lambda^3} + Q_{\min}^{1/2} \right] \left[1 + \frac{\mathcal{H}(\tau)}{\lambda^4} \right]^{-1/2}.$$

Now, if $\lambda \geq Q_{\min}^{-1/2} \mathcal{H}^{1/3}(\tau)$ and $\lambda \geq \mathcal{H}^{1/4}(\tau)$, then

$$(2.3.28) \quad d_{\lambda}(\tau) \leq 2Q_{\min}^{1/2}$$

and $d_{\lambda(\tau)}(\tau) \rightarrow 0$ as $\tau \rightarrow 0$.

We have once more analyzed here the estimate which is uniform with respect to λ . Analogously as before, we could derive an alternative, not uniform estimate.

Let us now analyze in detail the finite element method (2.2.5) for the problem P_{λ} with $q_n = 1$, $\tau = \tau_n = T/N$, $n = 1, 2, \dots, N$. We get now

$$(2.3.29) \quad y_1 \left(1 + \frac{\tau \lambda^2}{2} \right) = a\lambda + \frac{\lambda}{2} \int_0^{\tau} f(t)(\tau-t) dt$$

$$\frac{y_{n+1} - y_n}{\tau} + \frac{\lambda^2}{2} (y_n + y_{n+1}) = \lambda \int_{t_{n-1}}^{t_{n+1}} f(t) g_n(t) dt,$$

where

$$g_n(t) = \begin{cases} \frac{1}{\tau}(t - t_n) & \text{for } t_{n-1} < t < t_n \\ \frac{1}{\tau}(t_{n+1} - t) & \text{for } t_n < t < t_{n+1} \end{cases}$$

and

$$y_n = u_{\underline{q}}((n-\frac{1}{2})\tau) = u_{\underline{q}}|_{I_n}(t_n), \quad n = 1, 2, \dots, N.$$

We see that the finite element method is the well known Crank-Nicholson difference scheme. By Theorem 2.2.2 we have

$$(2.3.30) \quad \|u_0 - u_{\underline{q}}\|_X \leq CN \inf_{w \in S} \|u_0 - w\|_X,$$

where u_0 is the exact solution of the problem P_λ . Obviously, for smooth function as e.g., $u_0 = t^2$ we have

$$\inf_{w \in S} \|u_0 - w\|_X \geq CN^{-1}$$

and (2.3.30) suggest that the method does not converge.

On the other hand, by the classical finite difference approach we have

$$(2.3.31) \quad |u_0(t_n) - y_n| \leq \begin{cases} CN^{-1} \|u_0^{(2)}\|_{L_\infty(I)} \\ CN^{-2} \|u_0^{(3)}\|_{L_\infty(I)} \end{cases}.$$

Letting $Pu_0 \in S$ be such that $Pu_0|_{I_n} = u_0(t_n)$ we get

$$(2.3.32) \quad \|u_0 - u_g\|_X \leq \|u_0 - Pu_0\|_X + \|Pu_0 - u_g\|_X \leq CN^{-1} \|u_0^{(2)}\|_{L_\infty(I)}.$$

Hence we obtained the convergence in contrast to the estimate (2.3.30).

Let us first note that using (2.3.11) we get instead of (2.3.30), where C is independent of λ , the estimate

$$(2.3.33) \quad \|u_0 - u_g\|_X \leq C(1+\lambda^4) \inf_{w \in S} \|u_0 - w\|_X.$$

Further, we remark that the estimate (2.3.31) assumes high smoothness. Let, for example

$$u_0(t) = a\lambda e^{-\lambda^2 t}.$$

Then

$$\|u^{(k)}\|_{L_\infty(I)} = C\lambda^{2k+1}$$

and for $\lambda = N$ we get from (2.3.32)

$$\|u_0 - u_g\|_X \leq CN^{-1} \lambda^5 = CN^4,$$

while by (2.3.30) we get a bounded error.

3. The h-p version for the parabolic problem.

3.1. Preliminaries and problem formulation.

Let $\Omega \subset \mathbb{R}^2$ be a bounded, Lipschitz domain with a piecewise analytic boundary Γ . Let $D = I \times \Omega$, $I = (0, T)$. Then we will consider the problem

$$\begin{aligned} \frac{\partial u}{\partial t} - \Delta u &= f \quad \text{on } D \\ (3.1.1) \quad u &= 0 \quad \text{on } I \times \Gamma \\ u(0, x) &= g(x) \quad \text{on } \Omega. \end{aligned}$$

Let, as in [1], $X = X(D) = L_2(I, \dot{H}^1(\Omega))$ with

$$(3.1.2) \quad \|u\|_X^2 = \int_0^T \|u\|_{\dot{H}^1(\Omega)}^2 dt$$

and $Y = Y(D)$ be the completion of

$$\begin{aligned} \hat{C}^1 &= \{v \in C^\infty(\bar{I}) \mid v(t, x) \text{ has for any } t \in I \text{ compact support} \\ &\quad \text{in } \Omega \text{ and } v(T, x) = 0\} \end{aligned}$$

in the norm

$$(3.1.3) \quad \|v\|_Y^2 = \int_0^T (\|\dot{v}\|_{H^{-1}(\Omega)}^2 + \|v\|_{\dot{H}^1(\Omega)}^2) dt$$

where $\dot{v} = \frac{\partial v}{\partial t}$.

On $X \times Y$ we consider the bilinear form

$$(3.1.4) \quad B(u, v) = \int_0^T \int_{\Omega} (-u \dot{v} + \nabla u \nabla v) dx dt$$

and the problem P: Given $F \in Y'$, find $u_0 \in X$ such that

$$(3.1.5) \quad B(u, v) = F(v) \quad \forall v \in Y.$$

In [1] it has been shown that problem P has a unique solution for any

$F \in Y'$ and the solution u_0 of the problem P is a weak solution of (3.1.1) with

$$F(v) = \int_{\Omega} g(x)v(0,x)dx + \int_0^T \int_{\Omega} f(t,x)v(t,x)dxdt.$$

3.2. The semidiscrete problem. Discretization in x .

Let $R \subset \hat{H}^1(\Omega)$ be a finite dimensional subspace of functions and

$$S = \{u \in X \mid u(t,x) \in R \quad \forall t \in I\},$$

$$V = \{v \in Y \mid v(t,x) \in R \quad \forall t \in I\}.$$

The semidiscrete approximation of the solution for the problem P is defined as the function $u_s \in S$ satisfying

$$(3.2.1) \quad B(u_s, v) = F(v) \quad \forall v \in V.$$

The following estimate of the error for this semidiscretization has been derived in [1],

$$(3.2.2) \quad \|u_0 - u_s\|_X \leq C(\mathcal{H}(R))^{-1} \inf_{w \in S} \|u_0 - w\|_X,$$

under the assumption that the space R has the property \mathcal{H} , i.e., there is a number $\mathcal{H}(R) < 1$ such that

$$(3.2.3) \quad \|u\|_{H_R^{-1}(\Omega)} \geq \mathcal{H}(R) \|u\|_{H^{-1}(\Omega)}$$

holds for any $u \in R$ with

$$\|u\|_{H_R^{-1}(\Omega)} = \sup_{v \in R} \frac{|\int_{\Omega} uv dx|}{\|v\|_{\hat{H}^1(\Omega)}}.$$

Moreover, it has been proven in [1] without assuming the property \mathcal{K} , that

$$(3.2.4) \quad \|u_0 - u_s\|_X \leq C \|u_0 - P_0 u_0\|_X,$$

where P_0 denotes the L_2 -orthogonal projection of X onto R .

Assuming that R is (ξ, η, ν) -regular, i.e.,

$$\xi(R) = \sup_{u \in R} \frac{\|u\|_{\dot{H}^1(\Omega)}}{\|u\|_{L_2(\Omega)}} < +\infty$$

$$\eta(R) = \sup_{\substack{u \in \dot{H}^1(\Omega) \\ \|u\|_{\dot{H}^1(\Omega)} \leq 1}} \|u - P_1 u\|_{L_2(\Omega)} < +\infty$$

$$\nu(R) = \sup_{\substack{u \in \dot{H}^1(\Omega) \\ \|u\|_{\dot{H}^1(\Omega)} \leq 1}} \|u - P_0 u\|_{L_2(\Omega)} < +\infty,$$

where P_0 , respectively P_1 , denotes the L_2 -, respectively, H^1 -projection operator of $\dot{H}^1(\Omega)$ onto R , we have proved

$$(3.2.5) \quad \mathcal{K}(R) \geq (1 + \xi(R)\eta(R))^{-1}.$$

Further we have shown that for any $u \in \dot{H}^1(\Omega)$, $\|u\|_{\dot{H}^1(\Omega)} = 1$

$$(3.2.6) \quad \|u - P_0 u\|_{\dot{H}^1(\Omega)} \leq (1 + \eta(R) + \nu(R))\xi(R).$$

3.3. Numerical examples.

Let us consider the following problem:

$$(3.3.1) \quad \begin{aligned} \dot{u} - u'' &= 0 \quad \text{for } (t, x) \in D = (0, 1) \times (0, 1) \\ u'(t, 0) &= 0, \quad u(t, 1) = 0 \quad \text{for } t \in (0, 1) \end{aligned}$$

$$u(x,0) = g(x) \text{ for } x \in (0,1),$$

where $\dot{u} = \frac{\partial u}{\partial t}$, $u' = \frac{\partial u}{\partial x}$, $u'' = \frac{\partial^2 u}{\partial x^2}$. Due to the symmetry, the above problem is equivalent to the problem on $\tilde{D} = (0,1) \times (-1,1)$, homogeneous Dirichlet boundary conditions, and a symmetric solution.

We will address the cases $g_1(x) = 1 - x^2$ and $g_2(x) = 1 - x$. In the first case, the solution has singularity in the point $x = 1$, $t = 0$, while in the second case the singularity is located in the point $x = 0$, $t = 0$.

First, we consider the case of one single element in time and three different meshes in space with discretization in space only. The first mesh (a) consists of one element only. The mesh (b) and (c) is composed by two elements with the nodal point in $x = 0.05$ and $x = 0.95$, respectively, and the space S polynomials of degree p on each element. Obviously for g_1

$$\inf_{w \in S} \|u - w\|_X$$

is essentially the same for the meshes (a) and (b) and the asymptotic rate is the same because the refinement is at the place where there is no singularity. In the case of mesh (c) we expect essentially two phases, the first one when the rate is exponential (in p) and the second one for sufficient high p when the rate is algebraic, the same as for meshes (a) and (b) (for more see [5]).

On the other hand $\mathcal{H}(R)$ is different for the meshes (a) and (b) and $\mathcal{H}(R)$ is the same for the meshes (b) and (c). In fact, (3.2.5) yields that

$$\mathcal{H}(R) \geq \left(1 + p \frac{h_{\max}}{h_{\min}}\right)^{-1},$$

where h_{\max} , respectively h_{\min} is the length of the maximal, respectively minimal, element. Figure 3.3.1 shows that the effect of $\mathcal{H}(R)$ does not

appear in the computations.

For the mesh (a) we have two estimates (3.2.2) and (3.2.4). The estimate (3.2.4) leads to the rate $O(p^{-4.5+\epsilon})$ as shown in [1], while the estimate (3.2.2) gives the rate $O(p^{-4+\epsilon})$, $\epsilon > 0$, arbitrary. The fact that the meshes (a) and (b) give practically identical results shows that the factor $\mathcal{H}(R)$ does not influence the results. Further, the case of the mesh (c) shows exactly what had to be expected, namely a concave curve (which would straighten for higher p). The theoretical slope $p^{-4.5}$ based on 3.37a [1] is shown in Figure 3.3.1, too.

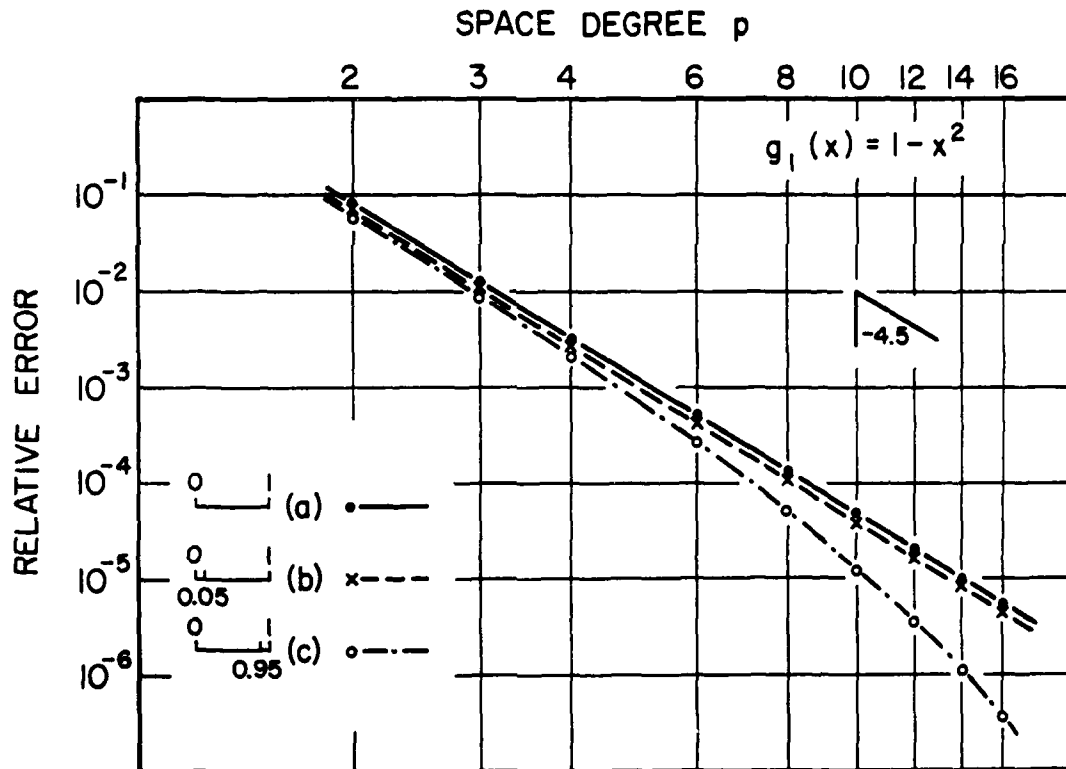


Figure 3.3.1. The relative error of the semidiscrete method (discretization in x) vs. the space degree p in the log log scale for the problem (3.3.1) with $g_1(x) = 1 - x^2$ (slope based on 3.37 a [1]).

- (a) one space element $h = 1$.
- (b) two space elements $h_1 = 0.05$, $h_2 = 0.95$
- (c) two space elements $h_1 = 0.95$, $h_2 = 0.05$.

Figure 3.3.2 shows the analogous results for $g_2(x) = 1 - x$. Now the singularity is in $x = 0, t = 0$ and hence the mesh (b) is better than mesh (c) in contrast with the previous case. The theoretical slope $p^{-2.5}$ is displayed in Figure 3.3.2, also.

Hence we can conclude that the performance of the method is not sensitive to the mesh (at least in the one-dimensional case).

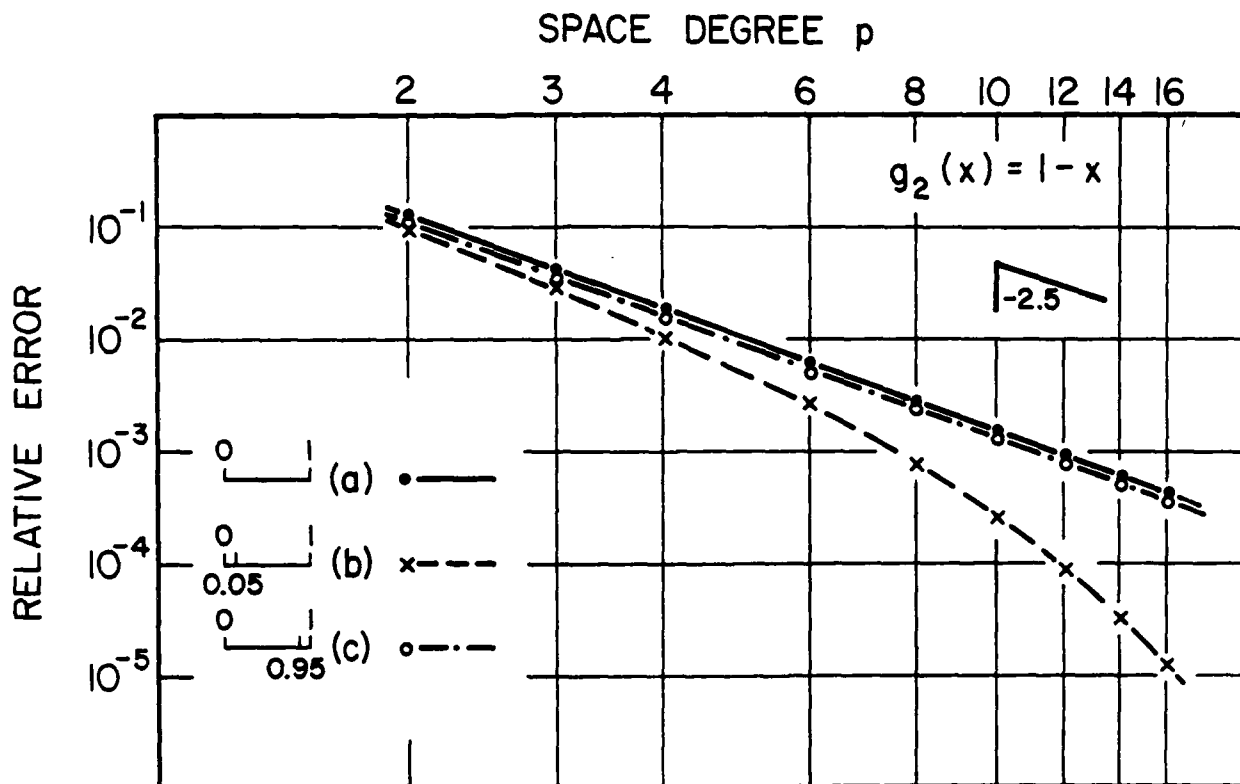


Figure 3.3.2. The relative error of the semidiscrete method (discretization in x) vs. the space degree p in the log log scale for the problem (3.3.1) with $g_2(x) = 1 - x$ (slope based on section 3.4 [1]).

- (a) one space element $h = 1$.
- (b) two space elements $h_1 = 0.05, h_2 = 0.95$
- (c) two space elements $h_1 = 0.95, h_2 = 0.05$.

3.4. The semidiscrete problem. Discretization in t .

Let now (as in Section 2.2)

$$(3.4.1) \quad \Delta_N : 0 = t_0 < t_1 < t_2 < \dots < t_N = T$$

be a subdivision of $(0, T)$ onto time intervals $I_n = (t_{n-1}, t_n)$ of the length $\tau_n = t_n - t_{n-1}$. Furthermore, let

$$(3.4.2) \quad \underline{q} = (q_1, q_2, \dots, q_N), \quad q_n \geq 1, \text{ an integer, } n = 1, 2, \dots, N$$

and

$$(3.4.3) \quad T_{\underline{q}-1} = \{u \in L^2(I) \mid u|_{I_n} \in S^{q_n-1}(I_n)\}$$

$$(3.4.4) \quad \hat{T}_{\underline{q}} = \{v \in C^0(\bar{I}) \mid v|_{I_n} \in S^{q_n}(I_n) \text{ and } v(T) = 0\}.$$

Define now

$$(3.4.5) \quad S_{\underline{q}} = T_{\underline{q}-1} \times \hat{H}^1(\Omega)$$

$$(3.4.6) \quad V_{\underline{q}} = \hat{T}_{\underline{q}} \times \hat{H}^1(\Omega).$$

For any $u \in S_{\underline{q}}$, respectively $v \in V_{\underline{q}}$, we have

$$(3.4.7) \quad u(t, x) = \sum_{i=1}^{\infty} \alpha_i(t) u_i(x),$$

where u_i are the eigenfunctions of the corresponding elliptic eigenvalue problem introduced in [1], with

$$\alpha_i \in T_{\underline{q}-1}$$

and

$$(3.4.8) \quad v(t, x) = \sum_{i=1}^{\infty} \beta_i(t) u_i(x),$$

with

$$\beta_1 \in \dot{f}_q^1.$$

Defining the bilinear form $B(u, v)$ on $S_q \times V_q$ by

$$(3.4.9) \quad B(u, v) = \int_0^T \left[\sum_{i=1}^{\infty} (-\alpha_i \dot{\beta}_i + \lambda_i^2 \alpha_i \beta_i) \right] dt,$$

where λ_i is the eigenvalue corresponding to u_i , $i = 1, 2, \dots$, we can use Theorem 2.2.1 to prove

Theorem 3.4.1. Let u_0 be the solution of the problem P and u_s its semi-discrete solution (discretization in t). Then, if $q_1 \geq q_2 \geq \dots \geq q_N \geq 1$ and $\tau_i/\tau_j \leq \sigma$ for $1 \leq i < j \leq N$ then

$$\|u_s - u_0\|_X \leq C q_1^{1/2N} \max(1, \sigma^{2q_1+(1/2)}) \inf_{w \in S_q} \|u_0 - w\|_X$$

3.5. Numerical examples.

Let us consider first the case when the initial function $g_3(x) = \cos \frac{\pi x}{2}$ and the semidiscrete (discretization in time) is used. In this case the problem essentially reduces to the ordinary differential equation because $g_3(x)$ is an eigenfunction. Figure 3.5.1 presents the results for different meshes. Theorem 3.4.1 gives the error estimate which strongly depends on σ . Figure 3.5.1 shows that the factor depending on σ is not essential, mainly because the eigenvalue λ is not too large and the major effect, as before, arises from the approximation. We mention that we could give for this case another estimate where the right hand side depends on λ .

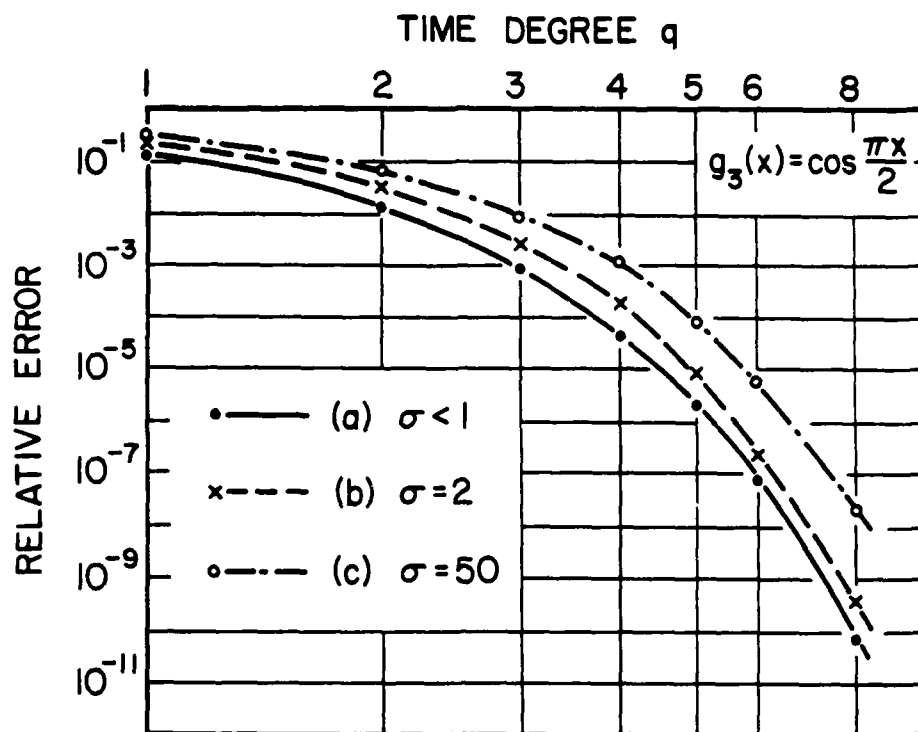


Figure 3.5.1. The relative error of the semidiscrete method (discretization in t) vs. the time degree q in the log log scale for the problem (3.1.1) with $g_3(x) = \cos \frac{\pi x}{2}$ and the following meshes:

- (a) $\tau_1 = 0.1, \tau_2 = 0.2, \tau_3 = 0.3, \tau_4 = 0.4, \sigma < 1$
- (b) $\tau_1 = 0.1, \tau_2 = 0.4, \tau_3 = 0.3, \tau_4 = 0.2, \sigma = 2$
- (c) $\tau_1 = 10/71, \tau_2 = 50/71, \tau_3 = 1/71, \tau_4 = 10/71, \sigma = 50.$

In the other example (Figure 3.5.2) we did not observe the dependence on σ (although for some solutions such dependence will occur, see (2.3.28)).

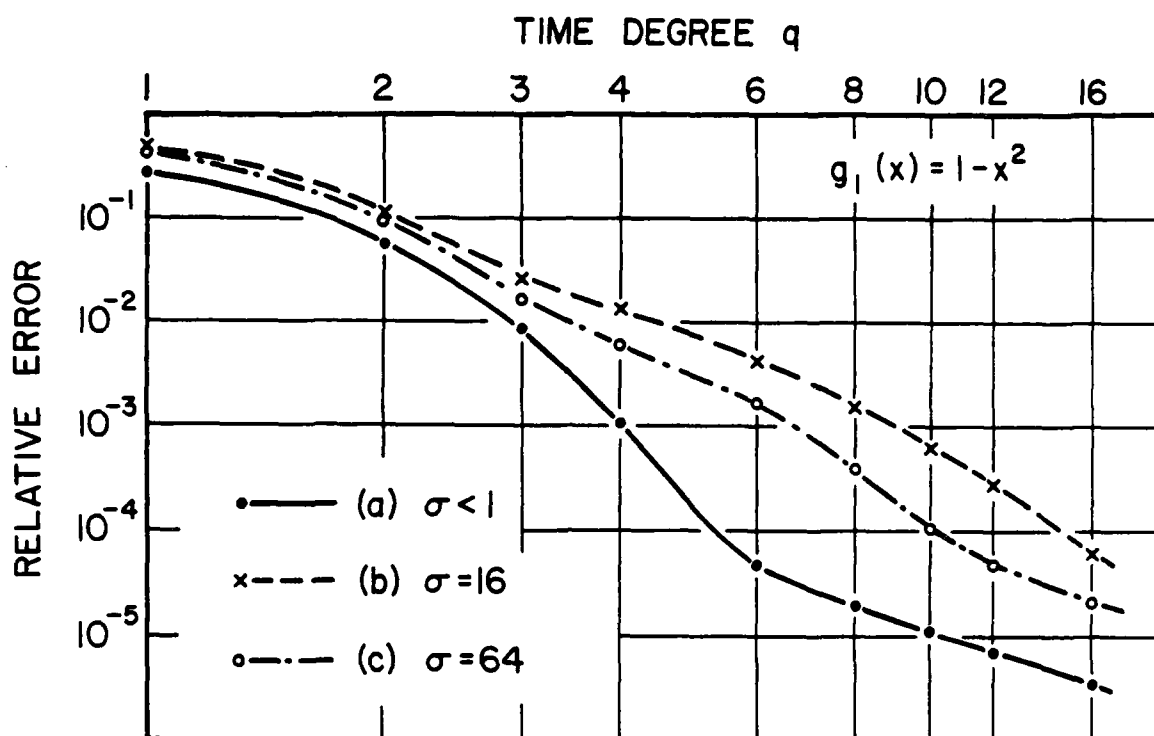


Figure 3.5.2. The relative error of the semidiscrete method (discretization in t) vs. the time degree q in the log log scale for the problem (3.1.1) with $g_1(x) = 1 - x^2$ and the following meshes:

- (a) $\tau_1 = 1/85, \tau_2 = 4/85, \tau_3 = 16/85, \tau_4 = 64/85, \sigma < 1$
- (b) $\tau_1 = 1/85, \tau_2 = 64/85, \tau_3 = 16/85, \tau_4 = 4/85, \sigma = 16$
- (c) $\tau_1 = 4/85, \tau_2 = 64/85, \tau_3 = 1/85, \tau_4 = 16/85, \sigma = 64.$

3.6. The complete discretization.

Let us define

$$(3.6.1) \quad S = S(R, q, \Delta) = T_{q-1} \times R$$

$$(3.6.2) \quad V = V(R, q, \Delta) = \hat{T}_q^1 \times R.$$

Combining the results of previous sections we get

Theorem 3.6.1. Let $u_0 \in X$ be the solution of the problem P and $u_s \in S$ be the finite element solution

$$(3.6.3) \quad B(u_s, v) = F(v) \quad \forall v \in V.$$

Then, if the space R has the property \mathcal{K} , $q_1 \geq q_2 \geq \dots \geq q_N \geq 1$, and $\tau_i/\tau_j \leq \sigma$ for $1 \leq i < j \leq N$, then

$$(3.6.4) \quad \|u_s - u_0\|_X \leq C q_1^{1/2} N(\mathcal{H}(R))^{-1} \max(1, \sigma^{2q_1+(1/2)}) \inf_{w \in S} \|u_0 - w\|_S$$

and

Theorem 3.6.3. Let $u_0 \in X$ be the solution of the problem P and $u_s \in S$ be the solution of (3.6.3). If ε_1 denotes the error of the semidiscrete method in (3.4.9) (discretization in t) and ε_2 is the error of the semidiscrete method (3.2.1) (discretization in x) and R is (ξ, η, ν) -regular, then

$$(3.6.5) \quad \|u_s - u_0\|_X \leq C(\varepsilon_1(1 + \eta(R) + \nu(R))\xi(R) + \varepsilon_2).$$

3.7. Numerical examples.

The h - p version of the finite element method gives large freedom to select elements in the space and time variables. This flexibility can be employed in various ways, for example, in connection with adaptive approaches. We also have seen that for $q = 1$, the method coincides with the Crank-Nicholson method and hence the h - p -version in t can be implemented as a special solver for the stiff ODE's arising from the solution of the semidiscrete method discussed in Section 3.2. Various aspects of this feature will be addressed in the forthcoming papers in this series. Here we will present some illustrative results related to the question of the optimal relation between p, q and a mesh in one dimensional setting we used earlier.

We will assume that in space (i.e., x -variable) only one element of

degree p is used, while in the time variable we use N elements of degree q . The shape functions are integrals of Legendre polynomials (in x), and in t we use Legendre polynomials as trial shape functions and integrals of Legendre polynomials and linear ones as test functions. Using the band solver we need

$$(3.7.1) \quad \bar{W} = N \cdot W$$

arithmetic operations, where

$$(3.7.2) \quad W \sim \begin{cases} p^3 & \text{for } 2q \geq p \\ 4q^3 p & \text{for } 2q < p. \end{cases}$$

We consider the problem (3.3.1) for $g(x) = g_2(x) = 1 - x$ and $g(x) = g_3(x) = \cos \frac{\pi x}{2}$ as the representatives of the solutions for unsmooth and smooth initial data problems.

In the case of $g(x) = g_2(x)$ we have used the radical mesh in t , i.e.

$$(3.7.3) \quad t_n = \left(\frac{n}{N} \right)^\gamma, \quad n = 0, 1, 2, \dots, N, \quad \text{with } \gamma = 3,$$

and for $g(x) = g_3(x)$ we have used the uniform mesh

$$(3.7.4) \quad t_n = \frac{n}{N}, \quad n = 0, 1, 2, \dots, N.$$

In the following figures we present the accuracy in dependence on (p, q) in the scale $\log(\text{relative error})$ vs. α with $q^\alpha = p$. We also show the work needed for computation based on (3.7.1 - 3.7.2).

For low q , the error is governed by the time integration, i.e., we deal here essentially with the case of semidiscrete method with time discretization (as discussed in section 3.4).

In contrast, high q 's essentially show the performance of the semidiscrete method with discretization in space only (as discussed in section 3.2).

In Figures 3.7.1 - 3.7.3 we present the results for $g(x) = g_2(x) = 1 - x$. We see that for any required accuracy the minimal computational work is needed when q is selected low and N large.

Note that the numbers in parentheses indicate the value of q used and e.g. the value 5.3E4 denotes that the approximate work $\bar{W} \approx 5.3 \cdot 10^4$ has been needed to compute the finite element solution.

In Figures 3.7.4 - 3.7.6, we show the analogous results for $g(x) = g_3(x) = \cos \frac{\pi x}{2}$. Here we see, in contrast to the previous case, that the best is to select one time element and large q .

In both cases we used the same p and q in all time intervals. The results show that the flexibility of the method, when properly employed, leads to the large increase of computational effectivity. Various adaptive approaches here will be very effective tools for such an optimal choice. These aspects will be addressed in the forthcoming paper.

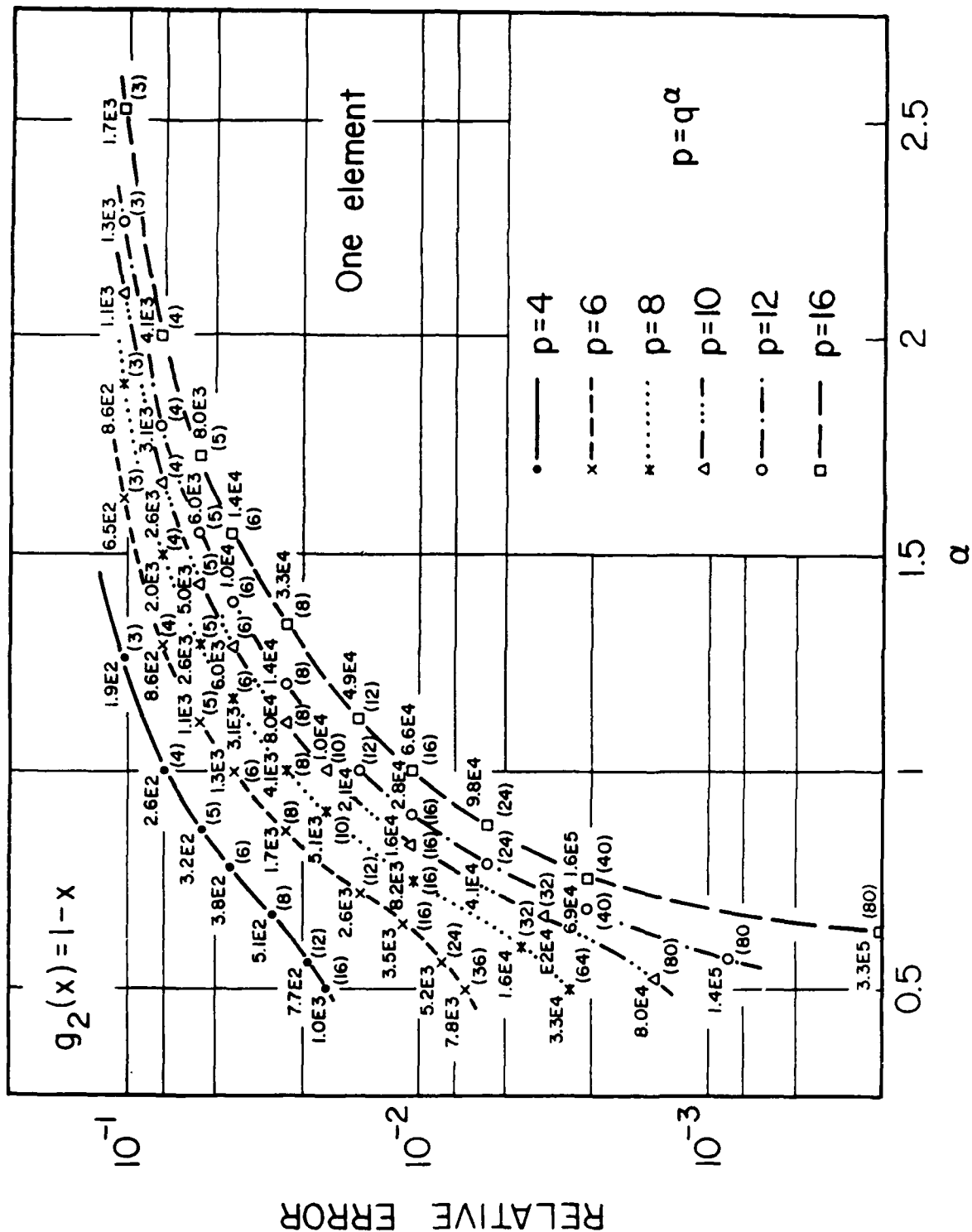


Figure 3.7.1. The performance of the p -version ($N = 1$) for the relationship $p = q^\alpha$. The initial function is $g_2(x) = 1 - x$.

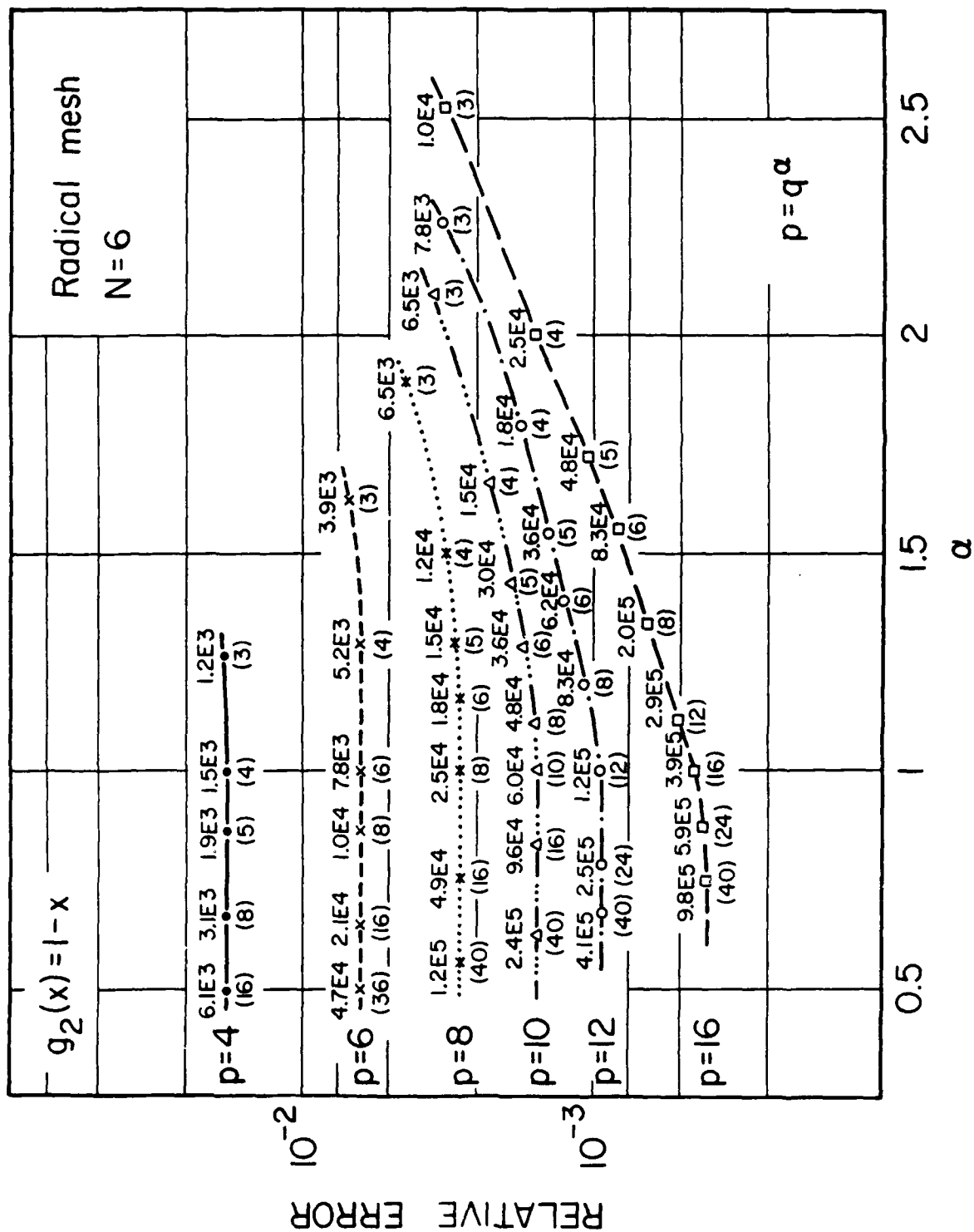


Figure 3.7.2. The performance of the h-p version ($N = 6$) for the relationship $p = q^\alpha$. The initial function is $g_2(x) = 1 - x$.

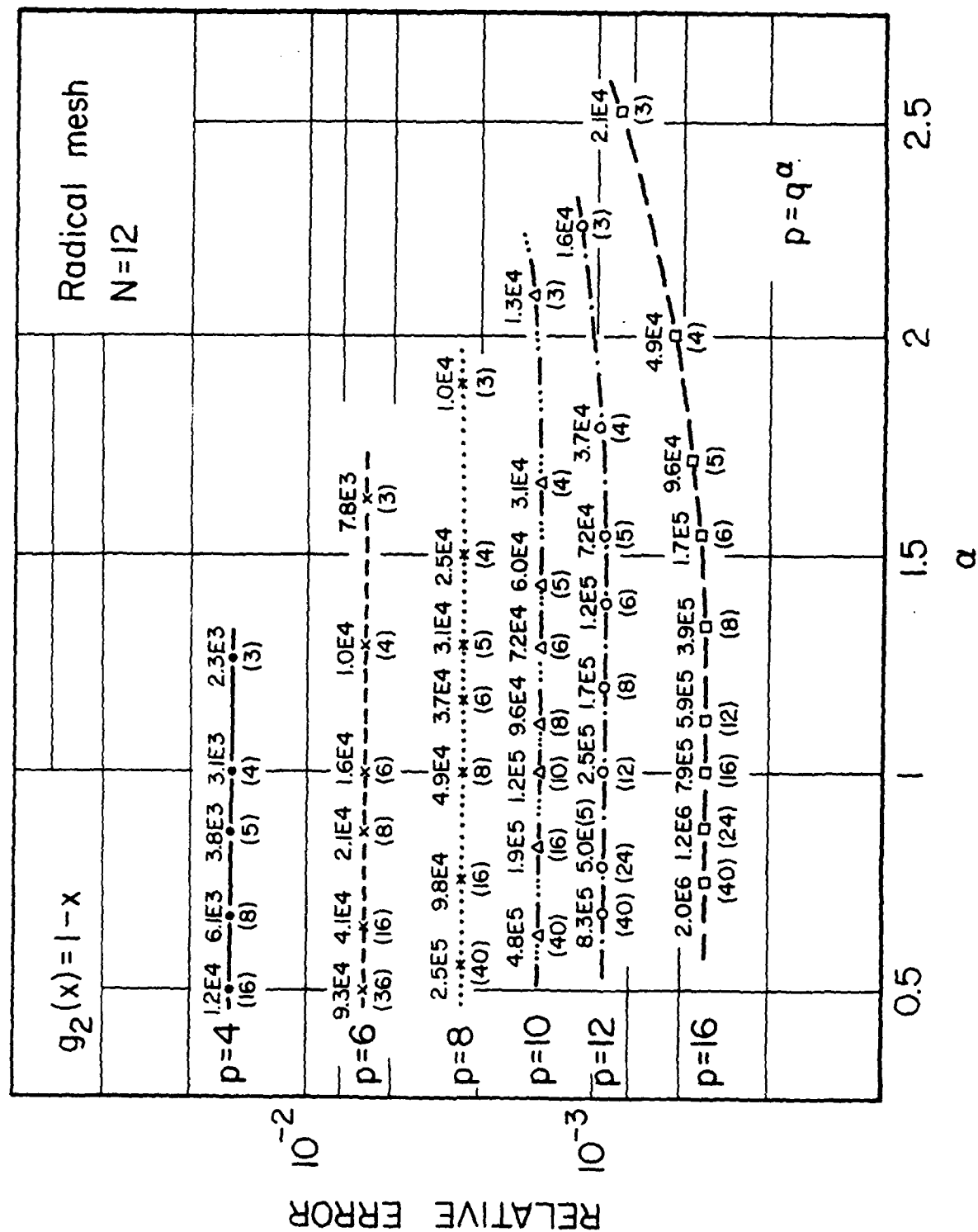


Figure 3.7.3. The performance of the h-p version ($N = 12$) for the relationship $p = q^\alpha$. The initial function is $g_2(x) = 1 - x$.

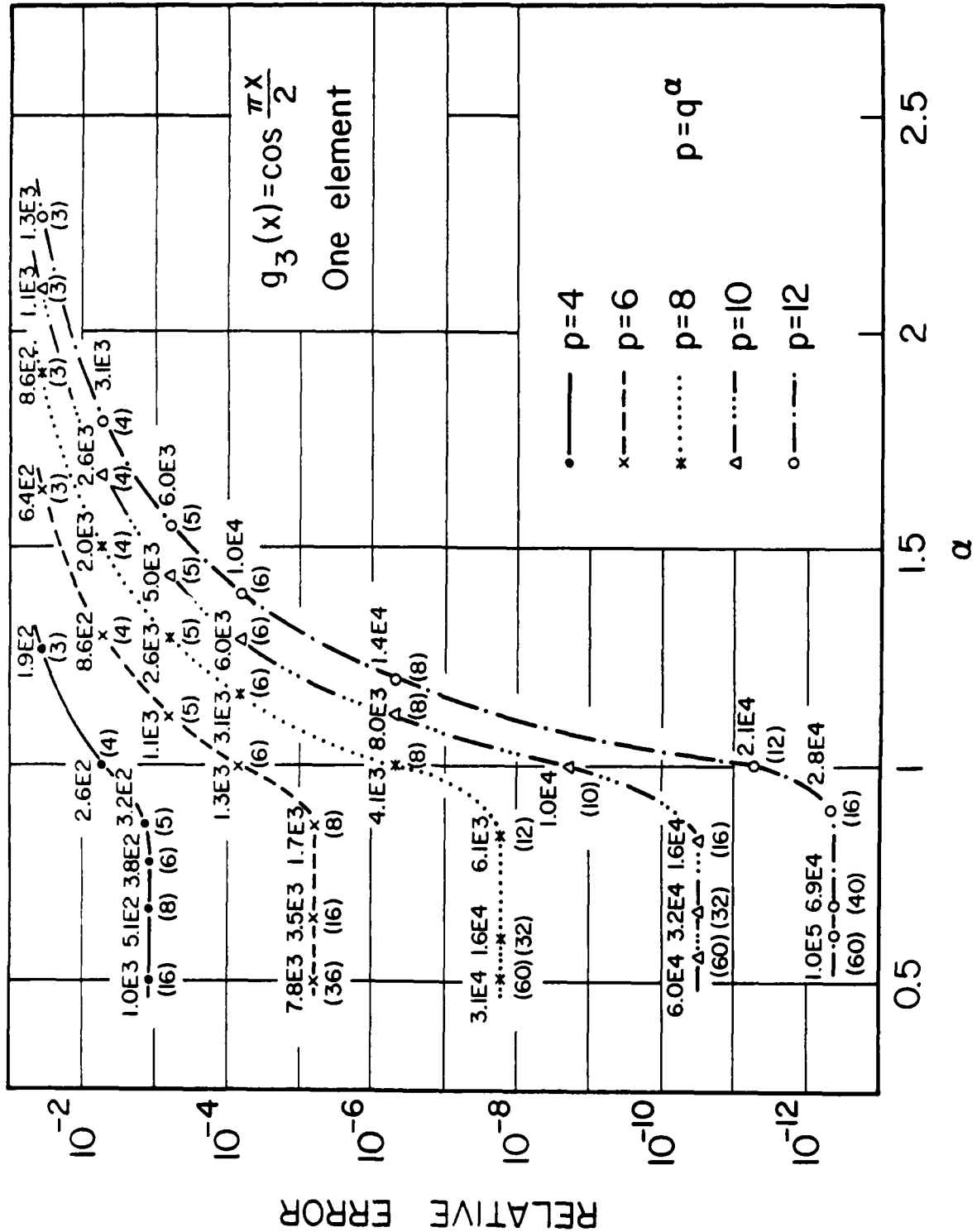


Figure 3.7.4. The performance of the p-version ($N = 1$) for the relationship $p = q^\alpha$. The initial function is $g_3(x) = \cos \frac{\pi x}{2}$.

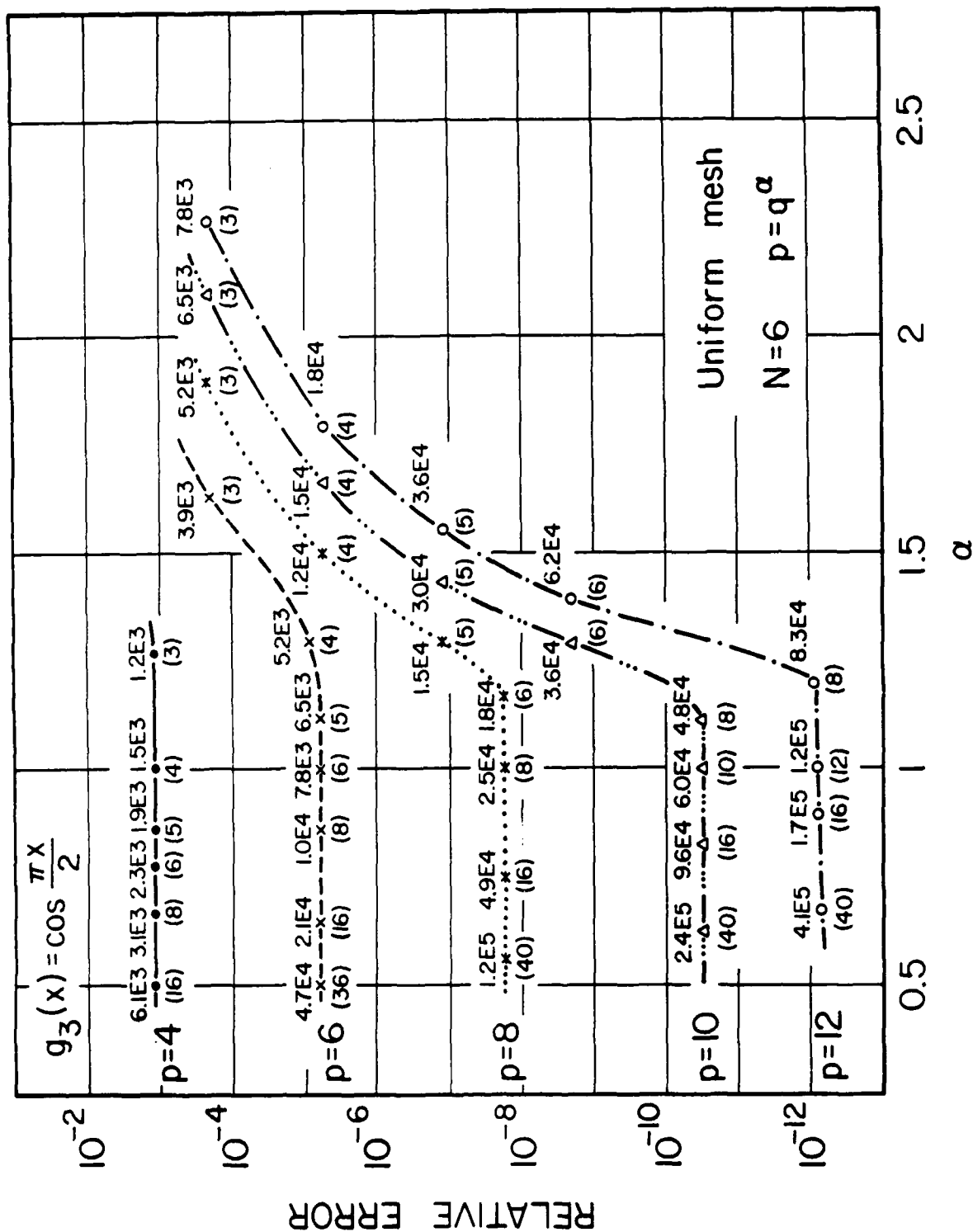


Figure 3.7.5. The performance of the h-p version ($N = 6$) for the relationship $p = q^\alpha$. The initial function is $g_3(x) = \cos \frac{\pi x}{2}$.

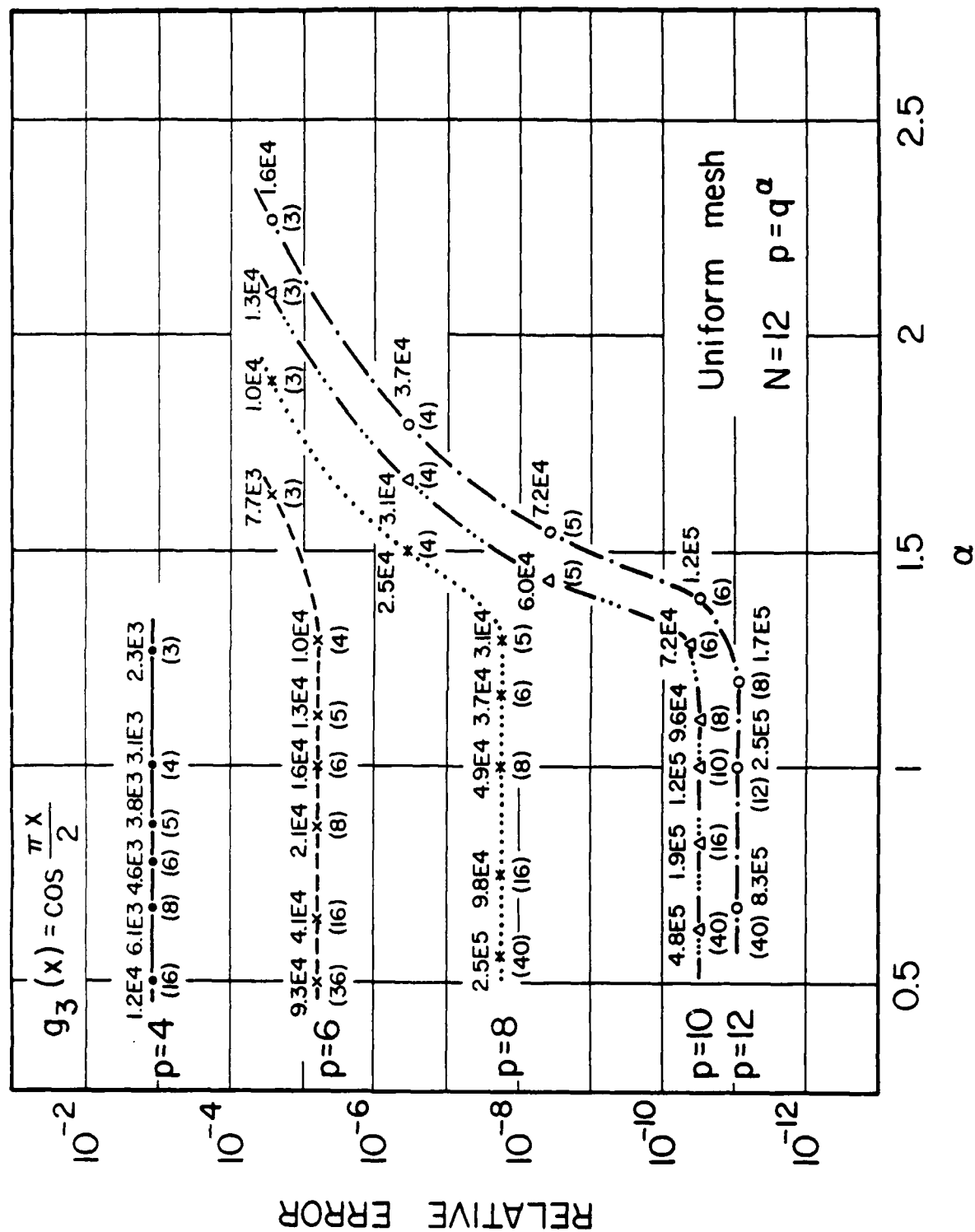


Figure 3.7.6. The performance of the h-p version (N = 12) for the relationship $p = q^\alpha$. The initial function is $g_3(x) = \cos \frac{\pi x}{2}$

References

- [1] Babuška, I., Janik, T., The h-p version of the finite element method for parabolic equations, Part I: The p-version in time, IPST, Univ. of Maryland, Technical Note BN-1077, July 1988 (to appear in Num. Meth. for Partial Diff. Eqs.).
- [2] Gradshteyn, I.S., Ryzhik, I.M., Table of Integrals, Series and Products, Academic Press, New York, London, 1965.
- [3] Babuška, I., Aziz, A.K., Survey Lectures on the Mathematical Foundations of the Finite Element Method, in The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, ed. A.K. Aziz, Academic Press, New York, London, 1972, 3-363.
- [4] Arnold, D.N., Babuška, I., Osborn, J., Finite Element Methods, Principles for Their Selection, Comp. Math. Appl. Mech. Engrg. 45 (1984), 57-96.
- [5] Gui, W., Babuška, I., The h,p and h-p version of the finite element method in one dimension, Part I, Part II, Numer. Math. 49 (1986), 577-612, 613-657.

The Laboratory for Numerical analysis is an integral part of the Institute for Physical Science and Technology of the University of Maryland, under the general administration of the Director, Institute for Physical Science and Technology. It has the following goals:

- o To conduct research in the mathematical theory and computational implementation of numerical analysis and related topics, with emphasis on the numerical treatment of linear and nonlinear differential equations and problems in linear and nonlinear algebra.
- o To help bridge gaps between computational directions in engineering, physics, etc., and those in the mathematical community.
- o To provide a limited consulting service in all areas of numerical mathematics to the University as a whole, and also to government agencies and industries in the State of Maryland and the Washington Metropolitan area.
- o To assist with the education of numerical analysts, especially at the postdoctoral level, in conjunction with the Interdisciplinary Applied Mathematics Program and the programs of the Mathematics and Computer Science Departments. This includes active collaboration with government agencies such as the National Bureau of Standards.
- o To be an international center of study and research for foreign students in numerical mathematics who are supported by foreign governments or exchange agencies (Fulbright, etc.)

Further information may be obtained from Professor I. Babuška, Chairman, Laboratory for Numerical Analysis, Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742.

END

FILMED

6-89

DTIC